Supplementary Material

Clinical parameters outperform molecular subtypes for predicting outcome in bladder cancer: Results from multiple cohorts including TCGA

Morera et al

Supplementary Material

Materials and methods

Patient cohort: Cohort-1: Specimens were obtained under a protocol approved by University of Miami's Institutional Review Board; individuals provided written consent. De-identified specimens and de-linked data were transferred to Augusta University under an approved protocol.

Oncomine-dataset: 151 high-grade MIBC specimens were from the Als, Lee and Sanchez-Carbayo-2 datasets and were accessed through Oncomine[1-3].

Surveillance, Epidemiology, and End Results Program (SEER): Patients were identified utilizing the bladder cancer site code (C67.0-C67.9) and International Classification of Diseases for Oncology codes (8120, 8122, 8123, 8130, 8131) for urothelial carcinoma. Excluding patients with no/unknown WHO/ISUP grade, the SEER-database has 73,354 patients with urothelial carcinoma (https://seer.cancer.gov/statfacts/html/urinb.html). Among 6,492 patients undergoing radical cystectomy, 5,564 had MIBC.

Molecular subtyping: For the Oncomine and TCGA datasets hierarchical clustering was performed using centroid linkage and Euclidean distance measurement[4]. Molecular subtyping models have been based on relative expression within a cohort[5-13]. The following method was used for assigning subtype identities by MCG-1 and MCG-Ext: Within a dataset the base expression levels of individual markers were different. For example, in the TCGA-dataset, median expression of FOXA1 was 12.1, while median level expression of KRT6C was 5.3. To control for the influence of the differences in base-level expression among markers, the median level of each marker in a dataset was used to determine whether a specimen had "high" or "low" expression of that specific marker. The high- and low-expression of BL markers was weighted as "+1" and "-1",

respectively. LU markers were designated in the opposite way, i.e. high expression: -1; low expression: +1. The sum of all BL and LU markers for each individual specimen was divided by the number of genes in a panel, resulting in a subtype score between "-1" and "+1" for each specimen. Therefore, tumors with a perfectly BL subtype expression pattern would have a score of +1, while tumors with a perfectly LU subtype expression pattern would have a score of -1; however, most tumors fell between +1 and -1. The subtype scores were z-normalized and the resulting z-scores were used to determine a specimen's subtype (BL, BL-like, LU-like, or LU).

Immunohistochemistry (IHC): IHC was performed using the IHC procedure described in detail previously[14]. The following antibodies were used and diluted in an antibody diluent solution from DAKO Thermo Fisher Scientific: KRT5/6 (Abcam ab17133), dilution: 1:250; KRT14 (EP1612Y, GeneTex GTX61595), dilution: 1:50; UPK2 (BC21, Biocare Medical ACI-3051-A), dilution: 1:100; UPK3 (Biovision 3620), dilution: 1:200; KRT20 (Ks20.8, Invitrogen MA5-13263), dilution: 1:100; FOXA1 (FOXA1/1241, GeneTex GTX34735), dilution: 0.5 ug/ml

Statistical Analysis: Analyses were performed using SAS9.4, JMP14.0 and GraphPad Prism software. We first evaluated the prognostic ability of the subtypes as a single parameter (univariate) to associate with clinical outcome. Since there were no multiple measurements over time, logistic regression model was used to univariately relate subtypes' function to clinical finding - i.e., metastasis, recurrence-free survival (RFS), cancer-specific survival (CSS) and overall-survival (OS). In an additional analysis, subtypes' association with time to metastasis, RFS, CSS or OS was determined while adjusting for all available demographic and clinical covariates described in Supplementary Table 1. We used the stepwise selection procedure to find the final Cox proportional hazard model which accurately described the outcome parameters (metastasis,

CSS, OS and RFS). Kaplan-Meier plots with log-rank statistics were prepared to determine if subtypes classified MIBC patients into risk categories for clinical outcome.

Figure S1

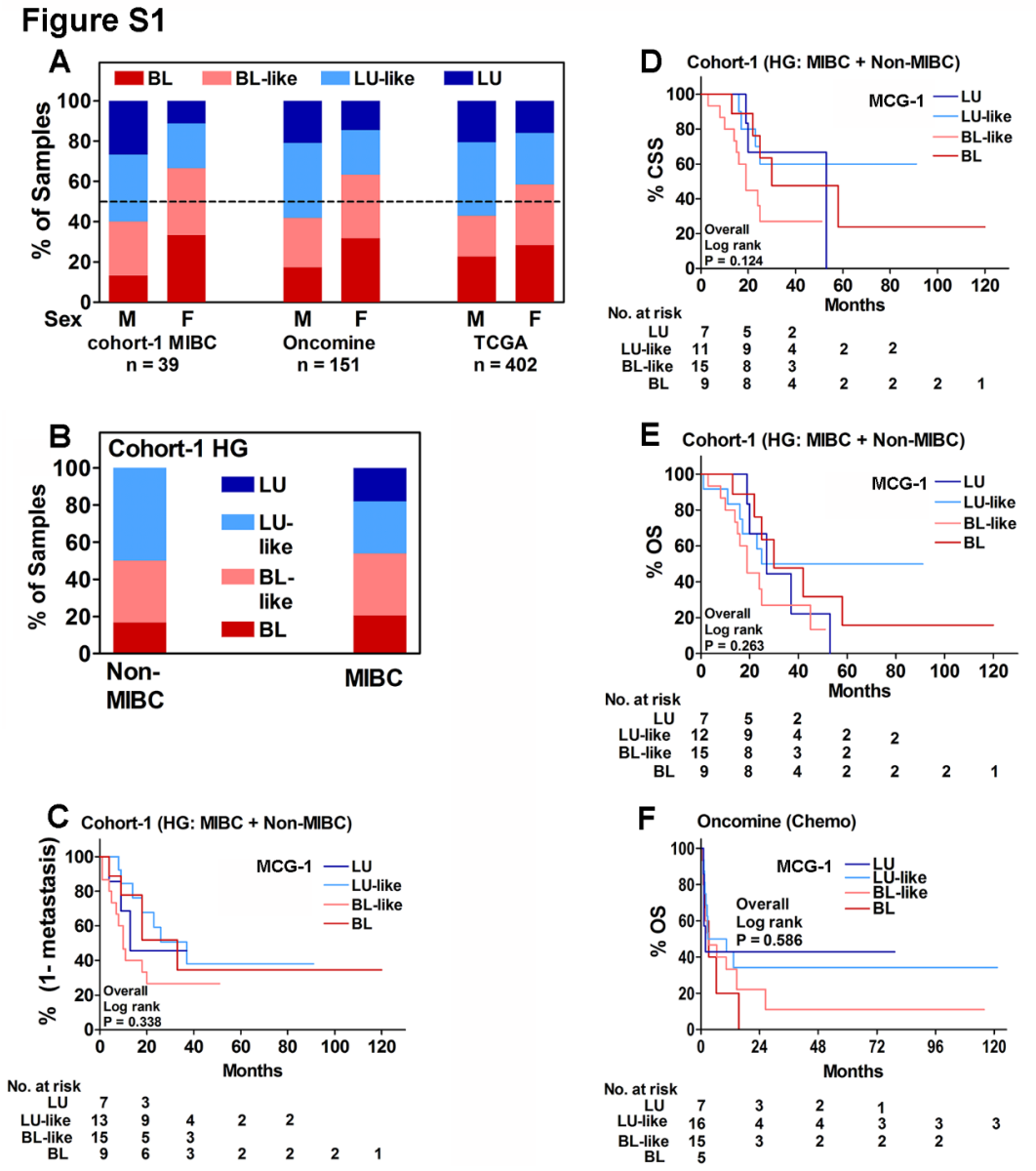

Figure S1: The distribution of sex among subtypes and the influence of stage and chemotherapy on the subtypes' association to clinical outcome. A. Distribution of molecular subtypes between male and female patients in the datasets based on MCG-1. **B.** The distribution of the molecular subtypes between high-grade (HG) non-MIBC and MIBC samples in cohort-1. **C – E.** Kaplan-Meier plots including non-MIBC specimens in cohort-1 for metastasis (B), CSS (C), and OS (D). **E.** Kaplan-Meier plot of OS for specimens in Oncomine-dataset treated with Cisplatin-based chemotherapy.

Supplementary Material

Table S1: Specimen and patient characteristics in TCGA, Oncomine and cohort-1 datasets. **Cohort-1** consists of 39 MIBC specimens in a cohort of 52 BC specimens (7 low-grade; 45 high-grade; 8 Ta; 5 T1; 39 MIBC). Five patients had concomitant carcinoma *in situ* (CIS). Note: Two patients died of BC but date of death was NA. **Oncomine dataset**[21-23]**: a: Als** dataset accessed from Oncomine™ in October 2017 contains 30 specimens. Of these, 29 were MIBC. Data available from Oncomine™ did not provide grade; however, in the study referenced by Oncomine™, Als *et. al.* reported that all patients had histology verified locally advanced or metastatic BC and therefore should be high-grade. **b: Lee Bladder** dataset was accessed from Oncomine™ in October 2017, and additional information was provided by Oncomine™ on August 2, 2018. The dataset contains 256 specimens. Sixty-eight specimens were normal bladder, 127 specimens were coded as non-MIBC (< T2) and 19 were coded as "low-grade" MIBC specimens[21]. After exclusion of these 214 specimens, 42 high-grade MIBC specimens remained. **c: Sanchez-Carbayo 2** dataset accessed from Oncomine™ in October 2017 contains 80 samples. All of the 80 samples are listed as MIBC specimens (G2: 6; G3: 74). **TCGA BC dataset**: When last confirmed in August 2018, the dataset has 436 samples. Twenty-three of these are "solid tissue normal" samples. One sample is from a "metastasis". Five samples are missing all RNAseq data. Five samples are non-muscle invasive (T0 N0 M0: 1; T1 N0 M0: 3; TX: 1). Exclusion of these 34 samples results in 402 MIBC specimens. **d: LN status:** lymph node invasion absent (-) or present (+). **e: LVI status:** lymphovascular invasion (LVI) absent (-) or present (+). **f**: TCGA and Oncomine™ datasets provide M-stage (synchronous metastasis) information. For cohort-1, metastasis data were available. **g - i**: In presenting the indicators (RFS, CSS, OS), indicator (+) means event, indicator (-) means no event. Therefore,

(+) designation indicates recurrence (i.e. progression) or death and the (-) designation indicates no recurrence (no progression) or survival. NA: Not available.

Supplementary Material

Patient characteristics					
Parameter	**Cohort-1**	**Oncomine™ -dataset**			**TCGA-dataset**
		Als[a]	Lee Bladder[b]	Sanchez-Carbayo-2[c]	
Specimens	52	29	42	80	402
Gender	Male: 40 Female: 12	Male: 22 Female: 7	Male: 33 Female: 9	Male: 55 Female: 25	Male: 296 Female: 106
Age (year)	Median: 67 Mean: 65.5 ± 9.8	Median: 60 Mean: 60.2 ± 7.3	Median: 69 Mean: 69 ± 9.3	Median: 69 Mean: 66.7 ± 9.6	Median: 68.5 Mean: 68.1 ± 10.6
Grade	High: 45 Low: 7	High: 29*	High: 42	G2: 6 G3: 74	Low: 21; High: 378; Missing: 3
T-stage	Non-MIBC: Ta: 8; T1: 5 MIBC: T2: 14; T3: 17; T4: 8	MIBC: T2: 2 T3: 13 T4: 14	MIBC: T2: 19 T3: 16 T4: 7	MIBC: T2: 11 T3: 57 T4: 12	MIBC: T2: 119 T3: 193 T4: 58 Missing: 32
N-stage[d]	(-): 30 (+): 17 Missing: 5	(-): 4 (+): 12 Missing: 13	(-): 30 (+): 11 Missing: 1	(-): 47 (+): 33	(-): 233 (+): 129 Missing: 40
LVI[e]	NA	NA	NA	NA	(-): 126 (+): 150 Missing: 126
M-stage or Metastasis[f]	(-): 25 (+) 26	(-): 15 (+): 14	(-): 39 (+): 3	NA	(-): 192 (+): 11 Missing: 199
Metastasis (Months)	25.3 ± 22.5	NA	NA	NA	NA
Adjuvant Cisplatin-based Chemotherapy	NA	Yes: 29 No: 0	Yes: 14 No: 14	NA	NA
RFS Indicator[g]	NA	NA	NA	NA	(-): 244 (+): 87 Missing: 71
RFS (months)	NA	NA	NA	NA	25.41 ± 27.9
CSS Indicator[h]	(-): 27 (+): 24	NA	(-): 20 (+): 22	(-): 35 (+): 45	NA
OS Indicator[i]	(-): 22 (+): 30	(-): 5 (+): 24	(-): 19 (+): 23	NA	(-): 222 (+): 176 Missing: 4
Survival (months)	30.5 ± 21.2	3.3 ± 3.6	29.9 ± 31.8	NA	26.5 ± 27.5

Supplementary Material

Table S2: Subtyping Panels

MCG-1	*BL: KRT5; KRT6A, KRT6B, KRT6C, KRT14* *LU: FOXA1, GATA3; UPK1B, UPK2, UPK3A; KRT20*
MCG-Ext	*BL: KRT5*, *KRT6A, KRT6B, KRT6C, KRT14*, CD44; EGFR; IL-6; TWIST1, SNAIL, *Vimentin,* CDH3, ZEB2; LU: *KRT20*, *FOXA1*, *GATA3*, *UPK1B, UPK2, UPK3A*, *CLDN3, CLDN4, CLDN7, CDH1, ERBB2, PPARG, FGFR3, CYP4B1, XBP1.*

References

1. Als AB, Dyrskjot L, von der Maase H, et al. Emmprin and survivin predict response and survival following cisplatin-containing chemotherapy in patients with advanced bladder cancer. *Clin Cancer Res.* 2007;13(15 Pt 1):4407-4414.
2. Sanchez-Carbayo M, Socci ND, Lozano J, Saint F, Cordon-Cardo C. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2006;24(5):778-789.
3. Kim WJ, Kim EJ, Kim SK, et al. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. *Mol Cancer.* 2010;9:3.
4. Babicki S, Arndt D, Marcu A, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* 2016;44(W1):W147-153.
5. Choi W, Porten S, Kim S, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell.* 2014;25(2):152-165.
6. Dadhania V, Zhang M, Zhang L, et al. Meta-Analysis of the Luminal and Basal Subtypes of Bladder Cancer and the Identification of Signature Immunohistochemical Markers for Clinical Use. *EBioMedicine.* Vol 122016:105-117.
7. Damrauer JS, Hoadley KA, Chism DD, et al. Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. *Proc Natl Acad Sci U S A.* Vol 1112014:3110-3115.
8. Ochoa AE, Choi W, Su X, et al. Specific micro-RNA expression patterns distinguish the basal and luminal subtypes of muscle-invasive bladder cancer. *Oncotarget.* 2016;7(49):80164-80174.
9. Rinaldetti S, Rempel E, Worst TS, et al. Subclassification, survival prediction and drug target analyses of chemotherapy-naive muscle-invasive bladder cancer with a molecular screening. *Oncotarget.* 2018;9(40):25935-25945.
10. Robertson AG, Kim J, Al-Ahmadie H, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell.* 2017;171(3):540-556.e525.
11. Seiler R, Ashab HAD, Erho N, et al. Impact of Molecular Subtypes in Muscle-invasive Bladder Cancer on Predicting Response and Survival after Neoadjuvant Chemotherapy. *European urology.* 2017;72(4):544-554.
12. Sjodahl G, Lauss M, Lovgren K, et al. A molecular taxonomy for urothelial carcinoma. *Clin Cancer Res.* 2012;18(12):3377-3386.
13. Seiler R, Gibb EA, Wang NQ, et al. Divergent Biological Response to Neoadjuvant Chemotherapy in Muscle-invasive Bladder Cancer. *Clin Cancer Res.* 2018.
14. Kramer MW, Escudero DO, Lokeshwar SD, et al. Association of hyaluronic acid family members (HAS1, HAS2, and HYAL-1) with bladder cancer diagnosis and prognosis. *Cancer.* 2011;117(6):1197-1209.